

# 爬蟲流程

## 尋找資料

- 技術基礎：  
urllib、requests、selenium
- 進階技術：  
多執行緒抓取、  
登入抓取、  
突破ip限制和伺服器抓取

## 請求

- 尋找消失的  
API/URL
- 溝通方式  
GET / POST
- 常用工具  
程式請求：  
urllib3/requests  
模擬請求：  
selenium

## 分析資料

- 工具語言：  
BeautifulSoup/  
PyQuery/JSON  
/Scrapy
- CSS選擇器：  
Selenium
- 路徑語言：  
Selenium/lxml

## 優化程式

- 模組化程式  
設計
- 定期定時的  
爬取觸發機  
制
- 異常處理設  
計